On the paper "A study on concave optimization via canonical dual function"

C. Zălinescu*

Abstract

In this short note we prove by a counter-example that Theorem 3.2 in the paper "A study on concave optimization via canonical dual function" by J. Zhu, S. Tao, D. Gao is false; moreover, we give a very short proof for Theorem 3.1 in the same paper.

Keywords: concave optimization, canonical dual function, counter-example

In [2] one says: "The primary goal of this paper is to study the global minimizers for the following concave optimization problem (primal problem (P) in short).

(P) min $P(x)$ (1.1)

s.t. $x \in D$,

where

$D = \{x \in R^n \mid \|x\| \leq 1\}$

and $P(x)$ is a smooth function in R^n and is strictly concave on the unit ball D, i.e. $\nabla^2 P(x) < 0$, on D."

Even if it is not said what is meant by "smooth function", from the context we think that P is assumed to be a C^2 function on R^n. One continues with "Let's consider the equation

$$\begin{cases} \nabla P(x) + \rho^* x = 0, \quad x^T x = 1, \\ \rho^* > 0. \end{cases} \qquad (2.1)$$

Suppose there are only finitely many of root pairs for (2.1):

$0 < \rho_1^* < \rho_2^* < \cdots < \rho_l^*$,

associated with feasible points on the unit sphere:

$\widehat{x}_1, \widehat{x}_2, \ldots, \widehat{x}_l$,

such that for each i,

$$\begin{matrix} \nabla P(\widehat{x}_i) + \rho_i^* \widehat{x}_i = 0, \quad \widehat{x}_i^T \widehat{x}_i = 1, \\ \rho_i^* > 0. \end{matrix} \qquad (2.2)\text{"}$$

Moreover, one says: "In Section 3, two sufficient conditions for determining a global minimizer are presented."

The results of [2] are the following.

"**Theorem 3.1.** If $\nabla^2 P(x) + \rho_l^* I > 0$ on $\|x\| \leq 1$, then $\widehat{x}_l$ is a global minimizer of (1.1)."

"**Theorem 3.2.** Suppose for $i = 1, 2, \ldots, l$, $\det\left[\nabla^2 P(\widehat{x}_i) + \rho_i^* I\right] \neq 0$ and $\frac{d^2 P_d(\rho_i^*)}{d\rho^{*2}} > 0$. Then $\widehat{x}_l$ is a global minimizer of (1.1)."

Related to these results we mention that Theorem 3.1 is (almost) trivial and Theorem 3.2 is false even for $n = 1$.

*University "Al.I.Cuza" Iaşi, Faculty of Mathematics, Romania, email: zalinesc@uaic.ro.

Indeed, because $\nabla^2 P(x)+\rho_l^* I > 0$ for $x \in D$, there exists $r > 1$ such that $\nabla^2 P(x)+\rho_l^* I > 0$ for $x \in D_r := \{u \in R^n \mid \|u\| < r\}$. Otherwise there exist the sequences $(x_k) \subset R^n$ with $1 < \|x_k\| \to 1$ and $(v_k) \subset S = \{u \in R^n \mid \|u\| = 1\}$ such that $v_k^T \left(\nabla^2 P(x_k) + \rho_l^* I\right) v_k \leq 0$ for every k. We may assume that $x_k \to x$ and $v_k \to v$; hence $x, v \in S$. It follows that $v^T \left(\nabla^2 P(x) + \rho_l^* I\right) v \leq 0$, contradicting our assumption. Since D_r is an open convex set we obtain that $P + \frac{1}{2}\rho_l^* \|\cdot\|^2$ is a (strictly) convex function on D_r. Because $\widehat{x}_l \in D \subset D_r$ and $\nabla\big(P+\frac{1}{2}\rho_l^* \|\cdot\|^2\big)(\widehat{x}_l) = \nabla P(\widehat{x}_l)+\rho_l^*\widehat{x}_l = 0$, we have that $\widehat{x}_l$ is a global minimizer of $P+\frac{1}{2}\rho_l^* \|\cdot\|^2$ on D_r. In particular we have that

$$P(\widehat{x}_l) + \tfrac{1}{2}\rho_l^* = P(\widehat{x}_l) + \tfrac{1}{2}\rho_l^* \|\widehat{x}_l\|^2 \leq P(x) + \tfrac{1}{2}\rho_l^* \|x\|^2 \leq P(x) + \tfrac{1}{2}\rho_l^* \quad \forall x \in D,$$

whence $P(\widehat{x}_l) \leq P(x)$ for every $x \in D$.

The proof above shows that whenever P is a C^2 function on an open set D_r containing D such that $\nabla^2 P(x) + \rho_l^* I > 0$ on D (or even less, $\nabla^2 P(x) + \rho_l^* I \geq 0$ on D_r) and $\overline{x} \in S$ and $\overline{\rho} \geq 0$ are such that $\nabla P(\overline{x}) + \overline{\rho}\,\overline{x} = 0$, then $\overline{x}$ is a global minimizer of P on D.

Related to [2, Th. 3.2], let us observe first that the condition $\frac{d^2 P_d(\rho_i^*)}{d\rho^{*2}} > 0$ is equivalent with $\widehat{x}_i^T \left[\nabla^2 P(\widehat{x}_i) + \rho_i^* I\right]^{-1} \widehat{x}_i < 0$.

Indeed, one says: "For $i = 1, 2, \ldots, l$, defined by

$\nabla P(\widehat{x}(\rho^*)) + \rho^*\widehat{x}(\rho^*) = 0, \quad \rho^* > 0, \;\; \widehat{x}(\rho_i^*) = \widehat{x}_i \qquad (2.3)$

a branch $\widehat{x}_i(\rho^*)$ is a continuously differentiable vector function on ρ^*." "In what follows, we suppress the index when focusing on a given branch according to the context.

The dual function [6] with respect to a given branch $\widehat{x}(\rho^*)$ is defined as

$P_d(\rho^*) = P(\widehat{x}(\rho^*)) + \frac{\rho^*}{2}\widehat{x}^T(\rho^*)\widehat{x}(\rho^*) - \frac{\rho^*}{2}. \qquad (2.6)$"

Note that [6] above is our reference [1].

In order to obtain a solution $\widehat{x}$ of (2.3) the authors use differential equations. In fact, let $F : R^n \times R \to R^n$ be defined by $F(x, \rho) := \nabla P(x) + \rho x$. Clearly, F is a C^1 function, $\nabla_\rho F(x, \rho) = x$, whence $\nabla_\rho F(\widehat{x}_i, \rho_i^*) = \widehat{x}_i \neq 0$. By the implicit function theorem a C^1 function $\widehat{x} : J \to R^n$ exists such that $F(\widehat{x}(\rho), \rho) = 0$ for $\rho \in J$ and $\widehat{x}(\rho_i^*) = \widehat{x}_i$, where J is an open interval containing ρ_i^*. It follows that

$$\nabla_x F(\widehat{x}(\rho), \rho)\widehat{x}'(\rho) + \nabla_\rho F(\widehat{x}(\rho), \rho) = \left[\nabla_x^2 P(\widehat{x}(\rho)) + \rho I\right] \widehat{x}'(\rho) + \widehat{x}(\rho) = 0 \quad \forall \rho \in J.$$

Because $\det\left[\nabla^2 P(\widehat{x}_i) + \rho_i^* I\right] \neq 0$, we may assume that $\det\left[\nabla^2 P(\widehat{x}(\rho)) + \rho_i^* I\right] \neq 0$ for all $\rho \in J$ (taking a smaller J if necessary). Hence

$$\widehat{x}'(\rho) = -\left[\nabla_x^2 P(\widehat{x}(\rho)) + \rho I\right]^{-1} \widehat{x}(\rho) \quad \forall \rho \in J.$$

From the expression of P_d in (2.6), using (2.3) we get

$$P_d'(\rho) = \nabla P(\widehat{x}(\rho))\widehat{x}'(\rho) + \tfrac{1}{2}\widehat{x}^T(\rho)\widehat{x}(\rho) + \rho\widehat{x}^T(\rho)\widehat{x}'(\rho) - \tfrac{1}{2} = \tfrac{1}{2}\widehat{x}^T(\rho)\widehat{x}(\rho) - \tfrac{1}{2},$$
$$P_d''(\rho) = \widehat{x}^T(\rho)\widehat{x}'(\rho)$$

for every $\rho \in J$. Using the expression of $\widehat{x}'(\rho)$ obtained above we get

$$P_d''(\rho_i^*) = -\widehat{x}^T(\rho_i^*) \left[\nabla_x^2 P(\widehat{x}(\rho_i^*)) + \rho_i^* I\right]^{-1} \widehat{x}(\rho_i^*) = -\widehat{x}_i^T \left[\nabla^2 P(\widehat{x}_i) + \rho_i^* I\right]^{-1} \widehat{x}_i.$$

This shows that instead of the condition $\frac{d^2 P_d(\rho_i^*)}{d\rho^{*2}} > 0$, which uses a quite complicated function, it was preferable to consider the condition

$$\widehat{x}_i^T \left[\nabla^2 P(\widehat{x}_i) + \rho_i^* I\right]^{-1} \widehat{x}_i < 0,$$

which is written using the data of the problem.

Example 1 *Consider* $P : R \to R$ *defined by* $P(x) = -x^4 - \frac{8}{5}x^3 - \frac{6}{5}x^2 + \frac{12}{5}x$. *Then*

$$P'(x) = -4x^3 - \tfrac{24}{5}x^2 - \tfrac{12}{5}x + \tfrac{12}{5}, \quad P''(x) = -12x^2 - \tfrac{48}{5}x - \tfrac{12}{5}.$$

We have that $P''(x) \le P''(-\frac{2}{5}) = -\frac{12}{25} < 0$ *for every* $x \in R$; *hence* P *is a strictly concave function. The system (2.1) becomes* $x = \pm 1$, $\rho^* = -x^{-1}P'(x)$, $\rho^* > 0$. *The solutions are* $(\widehat{x}_i, \rho_i^*)$, $i \in \{1, 2\}$, *where* $\widehat{x}_1 = -1$, $\widehat{x}_2 = 1$, $\rho_1^* = P'(-1) = 4$, $\rho_2^* = -P'(1) = \frac{44}{5}$. *Hence* $l = 2$ *and* $0 < \rho_1^* < \rho_2^*$. *The condition* $\widehat{x}_i^T \left[\nabla^2 P(\widehat{x}_i) + \rho_i^* I\right]^{-1} \widehat{x}_i < 0$ *becomes* $P''(\widehat{x}_i) + \rho_i^* < 0$ *for* $i \in \{1, 2\}$, *in which case* $\det\left[\nabla^2 P(\widehat{x}_i) + \rho_i^* I\right] \neq 0$. *But* $P''(-1) + 4 = -\frac{4}{5} < 0$, $P''(1) + \frac{44}{5} = -\frac{76}{5} < 0$. *Using [2, Th. 3.2] we obtain that* $\widehat{x}_2 = 1$ *is the global minimizer of* P *on* $[-1, 1]$. *However,* $P(-1) = -3 < -\frac{7}{5} = P(1)$, *proving that [2, Th. 3.2] is false.*

References

[1] D.Y. Gao, *Duality Principles in Nonconvex Systems, Theory, Methods and Applications*, Kluwer Academic Publishers (now Springer), Dordrecht, 2000.

[2] J. Zhu, S. Tao, D. Gao, *A study on concave optimization via canonical dual function*, Journal of Computational and Applied Mathematics 224 (2009), 459–464.